

Cluster consensus in discrete-time networks of multi-agents with adapted inputs

Yujuan Han, Wenlian Lu, *Member, IEEE*, Tianping Chen, *Senior Member, IEEE*

Abstract—In this paper, cluster consensus of multi-agent systems is studied via adaptive inputs. Here, we consider general graph topologies including time variation. The cluster consensus is defined by two aspects: the intra-cluster synchronization, that the state differences between agents in the same cluster converge to zero, and inter-cluster separation, that the states of the agents in different clusters are separated. For intra-cluster synchronization, the concepts and theories of consensus including the spanning trees, scramblingness, infinite stochastic matrix product and Hajnal inequality, are extended. With them, it is proved that if the graph has cluster spanning trees and all vertices self-linked, then static linear system can realize intra-cluster synchronization. For the time-varying coupling cases, it is proved that if there exists $T > 0$ such that the union graph across any T -length time interval has cluster spanning trees and all graphs has all vertices self-linked, then the time-varying linear system can also realize intra-cluster synchronization. Under the assumption of common inter-cluster influence, a sort of adapted inputs are utilized to realize inter-cluster separation, that each agent in the same cluster receives the same inputs and agents in different clusters have different inputs. In addition, the boundedness of the infinite sum of the inputs can guarantee the boundedness of the trajectory. Numerical examples are provided to illustrate the availability of our results.

Index Terms—Cluster Consensus, Multiagent System, Cooperative Control, Linear System

I. INTRODUCTION

In recent years, the multi-agent systems have broad applications [1], [2], [3]. In particular, the consensus problems of multi-agent systems have attracted increasing interests from many fields, such as physics, control engineering, and biology [4], [5], [6]. In network of agents, *consensus* means that all agents will converge to some common state. A consensus algorithm is an interaction rule how agents update their states. A large amount of papers concerning consensus algorithms have been published [7], [8], [9], [10], [11], most of which focused on the average principle, i.e., the current state of each agent is an average of the previous states of its own and its neighbors, which is implemented by communication between agents and can be described by the following difference equations for the discrete-time cases:

$$x_i(t+1) = \sum_{j=1}^n A_{ij} x_j(t), \quad i = 1, \dots, n, \quad (1)$$

where $x_i(t)$ denotes the state of agent i and $A = [A_{ij}]_{i,j=1}^n$ is a *stochastic matrix*. For a survey, we refer readers to [12]

and the references therein.

To realize consensus, the stability of the underlying dynamical system is crucial. Since the network can be regarded as a graph, the issues can be depicted by the graph theory. In the most existing literature, the concept of spanning tree is widely used to describe the communicability between agents in networks that can guarantee the consensus of (1). See [13], [14], [15].

It is widely known that the movement or/and defaults of the agents may lead the graph topology changing through time. So, it is inevitable to study the stability of the consensus algorithm in a time-varying environment, which can be described by the following time-varying linear system:

$$x_i(t+1) = \sum_{j=1}^n A_{ij}(t) x_j(t), \quad i = 1, \dots, n, \quad (2)$$

where each $A(t) = [A_{ij}(t)]_{i,j=1}^n$ is a stochastic matrix. There were a lot of literature, in which the stability analysis of (2) is investigated. Most of their results can be derived from the theories of infinite nonnegative matrix product and ergodicity of inhomogeneous Markov chain. Among them, the following should be highlighted. In [16], [17], the celebrated Hajnal's inequality was introduced and its generalized form was proposed in [18], to describe the compression of the differences between rows in a stochastic matrix when multiplied with another stochastic matrix that is scrambling. In [19], it was proved that a scrambling stochastic matrix could be obtained if a certain number of stochastic matrices that have spanning trees for their corresponding graphs were multiplied. So, in most of the papers involving stability analysis of (2), their sufficient conditions were expressed in the terms of spanning trees in the union graph across time intervals of a given length. See [8], [15] and the references therein. Besides, communication delays were also widely investigated [14], [20], [21] and nonlinear consensus algorithms were proposed [22].

All results mentioned above concern the consensus with a common consistent state. However, cluster consensus (synchronization) is considered to be more momentous in brain science [23] and engineering control [24], ecological science [25] and communication, engineering [26] and social science [27], and distributed computation [28]. This phenomenon is observed when the agents in networks are divided into several groups, called *clusters* in this paper, by the way that all individuals in the same cluster reach complete synchronization but the dynamics in different clusters do not coincide.

In this paper, we define the cluster consensus as follows. First, we divide the set of agents, denoted by \mathcal{V} , into disjoint

Y. Han, W. Lu, T. Chen are with School of Mathematical Sciences, Fudan University, No. 220 Handan Lu, Shanghai 200433, China.

W. Lu is also with the Max Planck Institute in Mathematics in the Sciences, Inselstr. 22, Leipzig D04103, Germany.

clusters, $\mathcal{C} = \{\mathcal{C}_1, \dots, \mathcal{C}_K\}$, with the properties:

- 1) $\mathcal{C}_p \cap \mathcal{C}_q = \emptyset$ for each $p \neq q$;
- 2) $\bigcup_{p=1}^K \mathcal{C}_p = \mathcal{V}$.

Second, letting $x(t) = [x_1(t), \dots, x_n(t)]^\top \in \mathbb{R}^n$ denote the state trajectory of all agents, of which $x_i(t)$ represents the state of $i \in \mathcal{V}$, we define *cluster-consensus* via the following aspects:

- 1) $x(t)$ is bounded;
- 2) We say that $x(t)$ *intra-cluster synchronizes* if $\lim_{t \rightarrow \infty} |x_i(t) - x_{i'}(t)| = 0$ for all $i, i' \in \mathcal{C}_p$ and $p = 1, \dots, K$;
- 3) We say that $x(t)$ *inter-cluster separates* if $\limsup_{t \rightarrow \infty} |x_i(t) - x_j(t)| > 0$ holds for each pair of $i \in \mathcal{C}_k$ and $j \in \mathcal{C}_l$ with $k \neq l$.

We say that a system reaches a *cluster consensus* if each solution $x(t)$ is bounded and satisfies the intra-cluster synchronization and inter-cluster separation, i.e., the items 1-3 are satisfied.

For this purpose, we introduce the following linear discrete-time system with external adapted inputs:

$$x_i(t+1) = \sum_{j=1}^n A_{ij} x_j(t) + I_i(t), \quad i \in \mathcal{C}_p, p = 1, \dots, K, \quad (3)$$

where $A = [A_{ij}]_{i,j=1}^n$ is a $n \times n$ stochastic matrix, $I_i(t)$ is an external scalar input and they are different with respect to clusters, which is used to realize inter-cluster separation. Also, we consider time-varying couplings that lead the following time-varying linear system with inputs:

$$x_i(t+1) = \sum_{j=1}^n A_{ij}(t) x_j(t) + I_i(t), \quad i \in \mathcal{C}_p, \\ p = 1, \dots, K. \quad (4)$$

Related Works. However, up till now, most papers in the literature mainly concern the global consensus. In some recent papers [29], [30], the authors addressed the cluster (group) consensus in networks with multi-agents and showed that (2) can reach cluster consensus if the graph topology is fixed and strongly connected and the number of clusters is equal to the period of agents. For continuous-time network with fixed topology, [29] proved that under certain protocol, the multi-agent network can achieve group consensus by discussing the eigenvalues and eigenvectors of the Laplacian matrix. [30] investigated group consensus in continuous-time network with switching topologies. However, all of these papers had a strong restriction in graph topologies and one important insight of cluster consensus: inter-cluster separation, has not been deeply investigated yet.

Our Contributions. In this paper, we derive sufficient conditions for cluster consensus in both (3) and (4). For this purpose, we extend the concepts and theories of graph and matrix such as spanning trees and scramblingness, as well as the transverse stability approach, to the cluster cases. For example, in the concept of “cluster spanning tree”, we release the concept of “spanning tree” as that each node in the same cluster has a common root vertex. Under the inter-cluster common influence condition, we construct the cluster-consensus subspace and transform the cluster consensus as

the stability of the cluster-consensus subspace. What’s more, we provide a novel and deep concept of cluster consensus by three aspects, as mentioned above: the boundedness of solutions, intra-cluster synchronization and inter-cluster separation. Therefore, the sufficient conditions we are looking for can divided into three parts. The boundedness of the finite sum of inputs guarantee that each solution is bounded. We extend the celebrated Hajnal inequality to the cluster case to prove that intra-cluster synchronization can be achieved if the union graph across any T -time interval has cluster spanning tree and all nodes self-linked for some integer T ; for inter-cluster separation, we prove that the trajectories in cluster-consensus subspace can be different for different clusters.

This paper is organized as follows. In section 2, we present some graph definitions and gives some notations required in this paper. In section 3, we firstly investigate the cluster consensus problem in discrete-time system with fixed topologies and present its cluster consensus criterion. Then we promote the criterion to the discrete-time system with switching topologies in section 4. Simulations are given in section 5 to verify the theoretical results. We conclude this paper in section 6.

II. PRELIMINARIES

In this section, we firstly recall some necessary notations and definitions that are related to graph and matrix theories and then generalized them into the cluster sense. We also present several lemmas and will be used later. For more details about the definitions, notations and propositions about the graph and matrix, we refer to textbooks [32], [33].

For a matrix A , denote A_{ij} the elements of A on the i th row and j th column. If the matrix A is denoted as the result of an expression, then we denote it by $[A]_{ij}$. A^\top denotes the transpose of A . For a set S with finite elements, $\#S$ denotes the number of elements in S . E denotes the identity matrix with a proper dimension. $\mathbf{1}$ denotes the column vector with all components equal to 1 with a proper dimension. $\rho(A)$ denotes the set of eigenvalues of a square matrix A . $\|z\|$ denotes a vector norm of a vector z and $\|A\|$ denotes the matrix norm of A induced by the vector norm $\|\cdot\|$.

An $n \times n$ matrix A is called a *stochastic matrix* if $A_{ij} \geq 0$ for all i, j , and $\sum_{j=1}^n A_{ij} = 1$ for $i = 1, \dots, n$. A stochastic matrix A is called *scrambling* if for any i and j , there exists k such that both A_{ik} and A_{jk} are positive.

A directed graph \mathcal{G} consists of a vertex set $\mathcal{V} = \{1, \dots, n\}$ and a directed edge set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$, i.e., an edge is an ordered pair of vertices in \mathcal{V} . \mathcal{N}_i denotes the neighborhood of the vertex v_i , i.e. $\mathcal{N}_i = \{j \in \mathcal{V} : (j, i) \in \mathcal{E}\}$. A (directed) *path* of length l from vertex v_i to v_j , denoted by $(v_{r_1}, v_{r_2}, \dots, v_{r_{l+1}})$, is a sequence of $l+1$ distinct vertices $v_{r_1}, \dots, v_{r_{l+1}}$ with $v_{r_1} = v_i$ and $v_{r_{l+1}} = v_j$ such that $(v_{r_k}, v_{r_{k+1}}) \in \mathcal{E}(\mathcal{G})$. The graph \mathcal{G} contains a *spanning (directed) tree* if there exists a vertex v_i such that for all the other vertices v_j there’s a directed path from v_i to v_j , and v_i is called the *root* vertex. Corresponding to the matrix scramblingness, we say that \mathcal{G} is scrambling if for any pair of vertices v_i and v_j there exists a common vertex k such that $(v_k, v_i) \in \mathcal{E}$ and $(v_k, v_j) \in \mathcal{E}$. We say that \mathcal{G} has self-links if $(v_i, v_i) \in \mathcal{E}$ for all $v_i \in \mathcal{V}$.

Ergodicity coefficient, $\mu(\cdot)$, was proposed to measure the scramblingness of a stochastic matrix. In [16], [17], the *Hajnal diameter* $\Delta(\cdot)$ was introduced to measure the difference of the rows in a stochastic matrix, and established his celebrated Hajnal's inequality $\Delta(AB) \leq (1 - \mu(A))\Delta(B)$, which indicated that the Hajnal diameter of stochastic matrix product AB strictly decreases w.r.t. B , if A is scrambling, i.e., $\mu(A) < 1$. The definitions of $\mu(\cdot)$ and $\Delta(\cdot)$ can be found in [16], [18].

An $n \times n$ nonnegative matrix A can be associated with a directed graph $\mathcal{G}(A) = \{\mathcal{V}, \mathcal{E}\}$ in such a way that $(v_j, v_i) \in \mathcal{E}$ if and only if $A_{ij} > 0$. With this correspondence, we also say A contains a spanning tree if $\mathcal{G}(A)$ contains a spanning tree. On the other hand, for a given graph \mathcal{G}_1 , we denote by $\mathcal{A}(\mathcal{G}_1) = \{A | \mathcal{G}(A) = \mathcal{G}_1\}$ the subset of stochastic matrices A such that $\mathcal{G}(A) = \mathcal{G}_1$.

For an infinite stochastic matrix sequence $\{A(t)\}_{t=1}^{\infty}$ with the same dimension, we use the following simplified symbol for a successive matrix product from t to s with $s > t$:

$$A_t^s \triangleq A(s)A(s-1) \cdots A(t).$$

For a constant matrix A , we denote its t -th power by A^t . [19] proved that if each stochastic matrix $A(t)$ has spanning trees and self-links, then A_t^s is scrambling if $s - t > n - 1$, where n is the dimension of the matrix $A(t)$ [34].

In this paper, we consider cluster dynamics of networks. First of all, for a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, we define a *clustering*, \mathcal{C} , as a disjoint division of the vertex set, namely, a sequence of subsets of \mathcal{V} , $\mathcal{C} = \{\mathcal{C}_1, \dots, \mathcal{C}_K\}$, that satisfies: (i) $\bigcup_{p=1}^K \mathcal{C}_p = \mathcal{V}$; (ii) $\mathcal{C}_k \cap \mathcal{C}_l = \emptyset$, $k \neq l$. Thus, we are able to extend the concepts of graph and matrix mentioned above to those in the cluster sense.

Definition 1: For a given clustering $\mathcal{C} = \{\mathcal{C}_1, \dots, \mathcal{C}_K\}$, we say that the graph \mathcal{G} has *cluster-spanning-trees* with respect to (w.r.t.) \mathcal{C} if for each cluster \mathcal{C}_p , $p = 1, \dots, K$, there exists a vertex $v_p \in \mathcal{V}$ such that there exist paths in \mathcal{G} from v_p to all vertices in \mathcal{C}_p . We denoted this vertex v_p as the root of the cluster \mathcal{C}_p .

It should be emphasized that the root vertex of \mathcal{C}_p and the vertices of the paths from the root to the vertices in \mathcal{C}_p do not necessarily belong to \mathcal{C}_p . It can be seen that the root vertex of a cluster does not necessarily same with other clusters. Therefore, the definition of the cluster-spanning-tree can be regarded as a generalization of that of spanning tree we mentioned above.

Definition 2: For a given clustering $\mathcal{C} = \{\mathcal{C}_1, \dots, \mathcal{C}_K\}$, we say that G is *cluster-scrambling* (w.r.t. \mathcal{C}) if for any pair of vertices $v_{p_1}, v_{p_2} \in \mathcal{C}_p$, there exists a vertex $v_k \in \mathcal{V}$, such that both (v_k, v_{p_1}) and (v_k, v_{p_2}) belong to \mathcal{E} .

Similarly, one can see that Definition 2 is a generalization of that of scramblingness we mentioned above. For a pair of vertices that are located in different clusters, they are not necessary to have a common linked vertex.

To measure the spanning-scramblingness, as a generalization from those in Hajnal [16], [17], we define the cluster ergodicity coefficient (w.r.t the clustering \mathcal{C}) of a stochastic

matrix A as

$$\mu_{\mathcal{C}}(A) = \min_{p=1, \dots, K} \min_{i, j \in \mathcal{C}_p} \sum_{k=1}^N \min(g_{ik}, g_{jk}).$$

It can be seen that $\mu_{\mathcal{C}}(A) \in [0, 1]$ and A is cluster-scrambling (w.r.t. \mathcal{C}) if and only if $\mu_{\mathcal{C}}(A) > 0$.

According to the definition of cluster consensus, we extend the definition of Hajnal diameter [16], [17], [18] to the cluster case:

Definition 3: For a matrix A , which has row vectors A_1, A_2, \dots, A_n and a given clustering \mathcal{C} , we define the cluster Hajnal diameter as

$$\Delta_{\mathcal{C}}(A) = \max_{p=1, \dots, K} \max_{i, j \in \mathcal{C}_p} \|A_i - A_j\|$$

for some norm $\|\cdot\|$.

It can be seen that $\Delta_{\mathcal{C}}(x) = 0$ is equivalent to the intra-cluster synchronization.

Similar to the results and the proof of Theorem 5.1 in [34], we can prove that the product of a sufficiently large number of stochastic matrices, all with cluster-spanning-trees, is cluster-scrambling.

Lemma 1: Suppose that each $A(t)$, $t = 1, \dots, n - 1$ is an n -dimensional stochastic matrix and has cluster-spanning-trees (w.r.t. \mathcal{C}) and self-links. Then the product A_1^{n-1} is cluster-scrambling (w.r.t. \mathcal{C}), i.e., $\mu_{\mathcal{C}}(A) > 0$.

Proof. For each cluster \mathcal{C}_p and each vertex $v_{p_1}, v_{p_2} \in \mathcal{C}_p$, let \mathcal{V}_1^t and \mathcal{V}_2^t be the vertices having directed links to v_{p_1} and v_{p_2} respectively via the graph $\mathcal{G}(A_1^t)$. The fact that each $A(t)$ has all nodes self-linked implies that the links in A_1^t will be reserved in A_1^{n-1} , if $n - 1 \geq t$, then the sets $\mathcal{V}_{1,2}^t$ increase respectively, that is, $\mathcal{V}_{1,2}^t \subset \mathcal{V}_{1,2}^{t+1}$ respectively. In the following, we are going to prove that the intersection between \mathcal{V}_1^{n-1} and \mathcal{V}_2^{n-1} is not an empty set by showing $\#\mathcal{V}_1^{n-1} + \#\mathcal{V}_2^{n-1} > n$. If so, we can complete the proof.

We prove it by reduction to absurdity. First, we suppose that $\mathcal{V}_1^t \cap \mathcal{V}_2^t = \emptyset$ for all $t \leq n - 1$. In the case of $t = 1$, since in $\mathcal{G}(A(1))$ there is a cluster root that has paths towards the vertices v_{p_1} and v_{p_2} , both $\mathcal{V}_{1,2}^1$ are not empty sets. Suppose that they do not have common vertices, which implies that $\#[\mathcal{V}_1^1 \cup \mathcal{V}_2^1] \geq 2$.

For any $t < n$, in the graph $\mathcal{G}(A(t+1))$ there exists a root vertex, denoted by v_1 , which has paths towards to v_{p_1} and v_{p_2} . We select their shortest paths: $(v_{k_1}, v_{k_2}, \dots, v_{k_P})$ and $(v_{l_1}, v_{l_2}, \dots, v_{l_Q})$, from v_1 to v_{p_1} and v_{p_2} respectively, with $v_{k_1} = v_{l_1} = v_1$, $v_{k_P} = v_{p_1}$ and $v_{l_Q} = v_{p_2}$. Under the assumption, at least one of the paths has at least one vertex not belonging to the corresponding \mathcal{V}_1^t or \mathcal{V}_2^t . Without loss of generality, we assume that $(v_{k_1}, v_{k_2}, \dots, v_{k_P})$ has vertices not belonging to \mathcal{V}_1^t and let v_{r_0} be the index such that

- for each $r > r_0$, $v_{k_r} \in \mathcal{V}_1^t$;
- $v_{k_{r_0}} \notin \mathcal{V}_1^t$.

This implies that

$$[A_1^{t+1}]_{v_{k_{r_0}}, v_{k_P}} \geq [A(t+1)]_{k_{r_0}, k_{r_0-1}} [A_1^t]_{v_{k_{r_0-1}}, v_{k_P}} > 0$$

holds. This implies that $v_{r_0} \in \mathcal{V}_1^{t+1}$. Hence,

$$\#(\mathcal{V}_1^{t+1} \cup \mathcal{V}_2^{t+1}) \geq \#(\mathcal{V}_1^t \cup \mathcal{V}_2^t) + 1 \geq t + 2.$$

That is, under the assumption of $\mathcal{V}_1^t \cap \mathcal{V}_2^t = \emptyset$ for all $t \leq n-1$, then $\#[\mathcal{V}_1^{t+1} \cup \mathcal{V}_2^{t+1}]$ strictly increases w.r.t. t . Hence, when $t = n-1$, $\#[\mathcal{V}_1^{t+1} \cup \mathcal{V}_2^{t+1}]$ will be greater than $n+1$, which implies that they should have a common vertex. This contradicts with the assumption, which completes the proof.

In [12], it has been proved that if a stochastic matrix A has spanning trees and all nodes self-linked, then the power matrix A^n converges to $\mathbf{1}\alpha$ for some row vector $\alpha \in \mathbb{R}^n$. Here, we conclude that the convergence can hold even without the spanning tree condition as a direct consequence from [33].

Lemma 2: If a stochastic matrix A has positive diagonals, then A^n is convergent exponentially.

III. CLUSTER CONSENSUS ANALYSIS OF DISCRETE-TIME NETWORK WITH STATIC COUPLING MATRIX

A. Invariance of the cluster consensus subspace

To seek sufficient conditions for cluster consensus, we firstly consider the situation that if the initial data $x(0) = [x_1(0), \dots, x_n(0)]^\top$ has already the cluster synchronizing structure, namely, $x_i(0) = x_j(0)$ for all $i, j \in \mathcal{C}_p$ with $p = 1, \dots, K$, then the cluster synchronization should be kept, i.e., $x_i(t) = x_j(t)$ for all $i, j \in \mathcal{C}_p$ with $p = 1, \dots, K$ and $t \geq 0$. In other words, the following subspace in \mathbb{R}^n w.r.t. the clustering \mathcal{C} :

$$\mathbb{S}_{\mathcal{C}} = \left\{ \begin{array}{l} x = [x_1, \dots, x_n]^\top \in \mathbb{R}^n : x_i = x_j \\ \text{for all } i, j \in \mathcal{C}_p \text{ with } p = 1, \dots, K \end{array} \right\},$$

named *cluster-consensus subspace*, is invariant through (3).

Definition 4: We say that the input $I(t)$ is *intra-cluster identical* if $I_i(t) = I_j(t)$ for all $i, j \in \mathcal{C}_p$ and all $p = 1, \dots, K$ and the stochastic matrix A has *inter-cluster common influence* if for each pair of p and p' , $\sum_{j \in \mathcal{C}_{p'}} a_{ij}$ is identical w.r.t. all $i \in \mathcal{C}_p$, in other words, $\sum_{j \in \mathcal{C}_{p'}} a_{ij}$ only depends on the cluster indices p and p' but is independent of the vertex $i \in \mathcal{C}_p$.

One can see that if two stochastic matrices A and B which have inter-cluster common influence w.r.t. the same clustering \mathcal{C} , then so it is with the product AB . In the following, similar to what we did in [35], we present the following lemma.

Lemma 3: If both the input is intra-cluster identical and the matrix A has inter-cluster common influence, then the cluster-consensus subspace is invariant through (3).

Proof. From the condition, we define

$$\beta_{p,p'} \triangleq \sum_{j \in \mathcal{C}_{p'}} a_{ij}$$

for any $i \in \mathcal{C}_p$ and $I_p(t) \triangleq I_i(t)$ for any $i \in \mathcal{C}_p$.

Assuming that $x(t) \in \mathbb{S}_{\mathcal{C}}$, we are to prove $x(t+1) \in \mathbb{S}_{\mathcal{C}}$, too. For this purpose, let $x_p(t)$ be the identical state of the cluster p at time t . Thus, for each \mathcal{C}_p and an arbitrary vertex $i \in \mathcal{C}_p$,

$$\begin{aligned} x_i(t+1) &= \sum_{p'=1}^K \sum_{j \in \mathcal{C}_{p'}} a_{ij} x_j(t) + I_i(t) \\ &= \sum_{p=1}^K \beta_{p,p'} x_{p'}(t) + I_p(t), \end{aligned}$$

which is identical w.r.t. all $i \in \mathcal{C}_p$. By induction, this completes the proof.

B. Intra-cluster synchronization

We assume a special sort of intra-cluster identical input as follows:

$$I_i(t) = \alpha_p u(t) \quad (5)$$

where $u(t)$ is a scalar function and $\alpha_1, \dots, \alpha_p$ are inter-different constants.

We extend the Hanjnal inequality [16], [17], [18] for measure the intra-cluster difference after multiplied with a stochastic matrix.

Lemma 4: Suppose that stochastic matrices A and B that both have the same dimension and inter-cluster common influence, then

$$\Delta_{\mathcal{C}}(AB) \leq (1 - \mu_{\mathcal{C}}(A)) \Delta_{\mathcal{C}}(B).$$

Proof. The idea of the proof is similar to that of the main result in [18]. Let

$$B = \begin{bmatrix} B_1 \\ \vdots \\ B_n \end{bmatrix}, \quad H = AB = \begin{bmatrix} H_1 \\ \vdots \\ H_n \end{bmatrix}$$

with $B_i = [B_{i1}, \dots, B_{in}]$ and $H_i = \sum_k a_{ik} B_k$, denoted by $[H_{i1}, \dots, H_{in}]$, for all $i = 1, \dots, n$.

For any pair of indices i and j both belonging to the same cluster \mathcal{C}_{p_0} , we have

$$H_i = \sum_{p=1}^K \sum_{k \in \mathcal{C}_p} a_{ik} B_k, \quad H_j = \sum_{p=1}^K \sum_{k \in \mathcal{C}_p} a_{jk} B_k.$$

Let $d_k = \min\{a_{ik}, a_{jk}\}$. Define a set of index vector of which different components are selected from different clusters:

$$W = \{w = [w_1, \dots, w_K] : w_p \in \mathcal{C}_p, p = 1, \dots, K\}.$$

For each $w \in W$, we define a class of convex combinations of B_1, \dots, B_n as follows:

$$G_w = \sum_{p=1}^K \left[\sum_{k \in \mathcal{C}_p, k \neq w_p} d_k B_k + (\beta_{p,p_0} - \sum_{k \in \mathcal{C}_p, k \neq w_p} d_k) B_{w_p} \right].$$

Since it can be seen that both H_i and H_j are in the convex hull of G_w for all $w \in W$,

$$\|H_i - H_j\| \leq \max_{w, w' \in W} \|G_w - G_{w'}\|.$$

In fact,

$$\begin{aligned} \|G_w - G_{w'}\| &\leq \sum_{p=1}^K (\beta_{p,p_0} - \sum_{k \in \mathcal{C}_p} d_k) \|B_{w_p} - B_{w'_p}\| \\ &\leq (1 - \mu_{\mathcal{C}}(A)) \Delta_{\mathcal{C}}(B). \end{aligned}$$

Hence,

$$\|H_i - H_j\| \leq (1 - \mu_{\mathcal{C}}(A)) \Delta_{\mathcal{C}}(B).$$

Therefore, $\Delta_{\mathcal{C}}(H) \leq (1 - \mu_{\mathcal{C}}(A)) \Delta_{\mathcal{C}}(B)$, which completes the proof due to the arbitrariness of $(i, j) \in \mathcal{C}_p$ and $p = 1, \dots, K$.

Remark 1: Lemma 4 indicates that if A has inter-cluster common influence, then the cluster-Hajnal diameter of Ax decreases. In addition, if A is cluster-scrambling, $d_C(Ax)$ is strictly less than $d_C(x)$.

Based on the previous lemmas, we give the following result concerning intra-cluster synchronization of (3).

Theorem 1: Suppose that both $u(t)$ and $\sum_{k=1}^t u(k)$ are bounded, $I(t)$ is defined by (5), and A is a stochastic matrix A with inter-cluster common influence, cluster-spanning trees and all diagonals positive. Then for any initial condition $x(0)$, (3) is bounded and intra-cluster synchronizes.

Proof. Let $x(k) = [x_1(k), \dots, x_n(k)]^\top$ be the solution of (3) that can be written as:

$$x(t+1) = A^{t+1}x(0) + \sum_{k=0}^t A^{t-k}I(k) \quad (6)$$

where $I(t) = \varsigma u(t)$ with $\varsigma = [\varsigma_1, \dots, \varsigma_n]^\top$ and

$$\varsigma_i = \alpha_p, \quad i \in \mathcal{C}_p. \quad (7)$$

There is some $Y > 0$ such that $|u(t)| \leq Y$, $|\sum_{k=0}^t u(k)| \leq Y$ hold for all $t \geq 0$.

By Lemma 2, we have $A^t = A^\infty + \epsilon(t)$, where $\|\epsilon(t)\|_\infty \leq M\lambda^t$ for some $M > 0$ and $\lambda \in (0, 1)$. Therefore,

$$\begin{aligned} \|x(t+1)\| &\leq \|A^{t+1}x(0)\| + \left\| \sum_{k=0}^t A^{t-k}\varsigma u(k) \right\| \\ &\leq \|x(0)\| + \|A^\infty\varsigma\| \left| \sum_{k=0}^t u(k) \right| + M \sum_{k=0}^t \lambda^{t-k} |u(k)| \\ &\leq \|x(0)\| + \|A^\infty\varsigma\| Y + MY \frac{1}{1-\lambda}. \end{aligned}$$

Therefore, the solution of system (3) is bounded.

By Lemma 1, we can find an integer N_1 such that for all $m \geq N_1$, A^m are cluster-scrambling. Denote $\eta = 1 - \mu(A^{N_1})$. For any t , let $t = pN_1 + l$ with some $0 \leq l < p$. We have

$$\Delta_C(A^{t+1}) \leq \eta^p \Delta_C(E_n)$$

which converges to zero as $t \rightarrow \infty$. In addition, since A^l has inter-cluster common influence and $\Delta_C(\varsigma) = 0$, then $\Delta_C(A^l\varsigma) = 0$ for all $l \geq 0$ can be concluded. Therefore, we have $\Delta_C(x(t+1)) \leq \Delta_C(A^{t+1}x(0))$ converges to zero as $t \rightarrow \infty$. This completes the proof.

C. Inter-cluster separation

Under the conditions of Theorem 1, the system can intra-cluster synchronize, namely, the states within the same cluster approach together. However, it is not known if the states in different clusters will approach together, too. A simple counterexample is that the matrix A with the inter-cluster common influence has (global) spanning trees and all diagonals positive and the inputs $\varsigma u(t)$ satisfies $\sum_{k=1}^\infty |u(k)|$ converges. In this case, the $u(t)$ converges to zero and the influence of the input to the system disappears. In this case, one can see that $x(t)$ reaches a global consensus, i.e., $\lim_{t \rightarrow \infty} a(t) = \mathbf{1}\alpha$ for some scalar α .

In this section, we investigate this problem by assuming that $u(t)$ is periodic with a period T and $\sum_{k=1}^T I(k) = 0$, which

guarantees that the sum of $u(t)$ is bounded. Construct a new matrix: $B = [\beta_{p,q}]_{p,q=1}^K$, where

$$\beta_{p,q} = \sum_{j \in \mathcal{C}_q} a_{ij}, \quad i \in \mathcal{C}_p \quad (8)$$

It can be seen that $\beta_{p,q}$ is independent of the selection of $i \in \mathcal{C}_p$.

Furthermore, we use the concept of “genericity” from the structural control theory [36], [37], [38] to investigate the inter-cluster separation. We define a set $\mathcal{T}(\mathcal{C}, \mathcal{G})$ w.r.t. a clustering \mathcal{C} and a graph \mathcal{G} , of which each element has form: $\{B, \tilde{\varsigma}, [u_1, \dots, u_{T-1}]\}$, where B is defined in (8) corresponding to the graph topology \mathcal{G} , $\tilde{\varsigma} \in \mathbb{R}^K$ is the vector to identify each cluster and defined as:

$$\tilde{\varsigma}_p = \alpha_p, \quad p = 1, \dots, K, \quad (9)$$

and $[u_1, u_2, \dots, u_{T-1}] \in \mathbb{R}^{T-1}$ is to define the periodic input:

$$\begin{aligned} u(\theta + kT) &= u_\theta, \quad \theta = 1, \dots, T-1, \\ \text{and } u(kT) &= -\sum_{j=1}^{T-1} u_j, \quad \forall k \geq 0. \end{aligned} \quad (10)$$

We can rewrite the system (3) as the following compact form:

$$x(t+1) = Ax(t) + \varsigma u(t), \quad (11)$$

where A and B has the relation as described by (8).

Definition 5: We say that for a given set $\mathcal{T}(\mathcal{C}, \mathcal{G})$ as defined above, (11) is *generically* inter-cluster separative (or cluster consensus) if for all most every triple $\{B, \tilde{\varsigma}, [u_1, \dots, u_{T-1}]\} \in \mathcal{T}(\mathcal{C}, \mathcal{G})$ and almost all initial $x(0) \in \mathbb{R}^n$, (11) can inter-cluster separate (or cluster consensus).

Before presenting a sufficient condition for generic inter-cluster separation, we give the following simple lemma.

Lemma 5: Suppose that the stochastic matrix A has the inter-cluster common influence. Then, for any pair of cluster \mathcal{C}_1 and \mathcal{C}_2 , either there are no links from \mathcal{C}_2 to \mathcal{C}_1 ; or for each vertex $v \in \mathcal{C}_1$, there are at least one link from \mathcal{C}_2 to v .

Theorem 2: Suppose that

- 1) every vertex in \mathcal{G} has a self-link;
- 2) \mathcal{G} satisfies the condition in Lemma 5 w.r.t \mathcal{C} ;
- 3) \mathcal{G} has cluster spanning trees.

Then ((11) reaches cluster consensus generically with respect to the set $\mathcal{T}(\mathcal{C}, \mathcal{G})$. In addition, the limiting consensus trajectories are periodic, that is, there exists some scalar periodic trajectories $v_p(t)$ with the period T for each cluster \mathcal{C}_p , $p = 1, \dots, K$, such that $\lim_{t \rightarrow \infty} |x_j(t) - v_p(t)| = 0$ if $j \in \mathcal{C}_p$.

Proof. We firstly prove the asymptotic periodicity. Recall the solution of (3) can be written as:

$$x(t+1) = A^{t+1}x(0) + \sum_{k=0}^t A^k \varsigma u(t-k). \quad (12)$$

By Lemma 2, one can see that A^n exponentially converges to A^∞ . Thus, we can find $M > 0$ and $\lambda \in (0, 1)$ such that

$\|A^t - A^\infty\| \leq M\lambda^t$. Let $Y = \max_{k=1, \dots, T} |u(k)|$. Thus, we have

$$\begin{aligned}
& \|x(t+lT+1) - x(t+1)\| \leq \|(A^{t+lT+1} - A^{t+1})x(0)\| \\
& + \left\| \sum_{k=0}^t A^k \varsigma [u(t+lT-k) - u(t-k)] \right\| \\
& + \left\| \sum_{k=t+1}^{t+lT} A^k \varsigma u(t+lT-k) \right\| \\
& = \|(A^{t+lT+1} - A^{t+1})x(0)\| + \left\| \sum_{k=t+1}^{t+lT} A^\infty \varsigma u(t-k) \right\| \\
& + \left\| \sum_{k=t+1}^{t+lT} [A^k - A^\infty] \varsigma u(t-k) \right\| \\
& \leq 2M\lambda^t \|x(0)\| + MY \|\varsigma\| \sum_{k=t}^{t+lT} \lambda^k \\
& = \left[2M\|x(0)\| + MY \|\varsigma\| \frac{1}{1-\lambda} \right] \lambda^t
\end{aligned}$$

for all l . Letting $t = mT + \theta - 1$ for any $m \in \mathbb{N}$ and $\theta = 1, \dots, T$, we have

$$\|x((m+l)T + \theta) - x(mT + \theta)\| \leq M_1 \lambda^{mT}$$

for some $M_1 > 0$. According to the Cauchy convergence principle, each $x(\theta + kT)$, $\theta = 1, \dots, T$, converges to some value as $k \rightarrow \infty$ exponentially, which implies that there exist T -periodic functions $v_p(t)$, $p = 1, \dots, K$, such that $|x_j(t) - v_p(t)| \rightarrow 0$ exponentially, if $j \in \mathcal{C}_p$.

Now, we will prove the consensus states in different clusters can be different generically.

Since each cluster synchronizes, we can pick a single vertex state from each cluster to represent the whole state of this cluster by the following way. We can divide the space \mathbb{R}^n into the direct sum of two subspaces: $\mathbb{R}^n = V_1 \oplus V_2$, where V_1 denotes the right eigenspace of A corresponding to the eigenvalue 1 and V_2 the right eigenspace of A corresponding to all other eigenvalues. Since all diagonals of A are positive, then the direct sum works and $AV_i \subset V_i$ holds for $i = 1, 2$. In addition, since the column vectors in A^∞ converges \mathcal{S}_C , $V_1 \subset \mathcal{S}_C$. So, $\mathbb{R}^n = \mathcal{S}_C + V_2$.

For any initial data $x(0) \in \mathbb{R}^n$, we can find $y^0 \in \mathcal{S}_C$ with the decompose $x(0) = y^0 + x(0) - y^0$ such that $x(0) - y^0 \in V$. Consider the following system restricted to \mathcal{S}_C :

$$y(t+1) = Ay(t) + I(t), \quad y(0) = y_0.$$

where $y(t) \in \mathcal{S}_C$ for all t .

Let $\delta x(t) = x(t) - y(t) \in V_2$. We have

$$\delta x(t+1) = A\delta x(t).$$

Theorem 1 tells that $x(t)$ converges to \mathcal{S}_C , which implies that $\lim_{t \rightarrow \infty} \delta x(t) = 0$, due to the direct sum property of \mathcal{S}_C and V . So, we can use $y(t)$ that belongs to \mathcal{S}_C to represent $x(t)$.

Since each component of y in the same cluster is identical, we can pick a single component from each cluster to lower-dimensional column vector $\tilde{y} \in \mathbb{R}^K$ with $\tilde{y}_p = y_i$ for some $i \in$

\mathcal{C}_p . Because of the exponential intra-cluster synchronization, we can write the equation for $y(t)$ as follows:

$$\tilde{y}(t+1) = B\tilde{y}(t) + \tilde{\varsigma}u(t) \quad (13)$$

where B is defined in (8) and $\varsigma = [\alpha_1, \dots, \alpha_K]^\top$. The inter-cluster separation can be equivalent to investigate the separation between components of \tilde{y} . One can see that for almost every B , B has K distinguishing left eigenvectors, denoted by ϕ_1, \dots, ϕ_K , corresponding to eigenvalues ν_1, \dots, ν_K (possibly overlapping). So, for almost every B with K left eigenvectors, let us write down the solution (13) at time nT as follows:

$$\begin{aligned}
\tilde{y}(nT+1) &= B^{nT+1}\tilde{y}(0) + \sum_{k=0}^{nT} B^{nT-k}\tilde{\varsigma}u(k) \\
&\rightarrow Z_1\tilde{y}(0) + Z_2\tilde{\varsigma}, \text{ as } n \rightarrow \infty,
\end{aligned}$$

where

$$Z_1 = \lim_{n \rightarrow \infty} B^{nT+1}, \quad Z_2 = \lim_{n \rightarrow \infty} \left[\sum_{k=0}^{nT} B^{nT-k}u(k) \right].$$

From Lemma 2, Z_1 does exists. Combined with $\sum_{k=0}^{nT-1} u(k) = 0$, we can conclude that the limit of Z_2 exists, too.

For an arbitrary fixed pair of (p, q) , with $p, q = 1, \dots, K$ and $p \neq q$, we are to show Z_2 can generically have different p -th and q -th components. In fact, for each k_1 with $|\nu_{k_1}| < 1$, noting that its associated left eigenvector is ϕ_{k_1} , we have

$$\begin{aligned}
& \phi_{k_1} \sum_{k=0}^{nT} B^{nT-k}u(k) \\
&= \phi_{k_1} \sum_{k=0}^{T-1} u(k) \sum_{z=0}^{n-1} \nu_{k_1}^{zT+k} + \phi_{k_1} \nu_{k_1}^{nT} u(0) \\
&\rightarrow \phi_{k_1} \frac{\sum_{k=0}^{T-1} u(k) \nu_{k_1}^k}{1 - \nu_{k_1}^T}, \text{ as } n \rightarrow \infty.
\end{aligned}$$

For each k_2 with $|\nu_{k_2}| = 1$, noting its associated left-eigenvector is ϕ_{k_2} , according to the fact that all diagonals in B is positive, from [33], we have $\nu_{k_2} = 1$ indeed. Then, we have

$$\phi_{k_2} \sum_{k=0}^{nT} B^{nT-k}u(k) = \phi_{k_2} \sum_{k=0}^{nT} u(k) = \phi_{k_2} u(0) = \phi_{k_2} u_T.$$

So, for almost $[u_1, \dots, u_{T-1}] \in \mathbb{R}^{T-1}$, the eigenvectors of Z_2 are the same with B and the corresponding eigenvalues are u_T and $\sum_{k=0}^{T-1} u(k) \nu_p^k / (1 - \nu_p^T)$. For almost every realization of $[u_i]_{i=1}^{T-1}$ and B , none of them is zero, which implies that Z_2 is nonsingular. That means that it is impossible for each pair of its rows to be identical. So, for almost all $\tilde{\varsigma}$ the p -th and q -th component of Z_2 are not identical. So, for almost every $\tilde{\varsigma}$, $Z_2\tilde{\varsigma}$ has no a pair of components identical. Therefore, we conclude that for almost every x_0 , associated with almost every $\tilde{y}(0)$, each pair of components in $Z_1\tilde{y}(0) + Z_2\tilde{\varsigma}$ are not identical.

We can arbitrarily select the cluster pair (p, q) and the exception cases of the statements above are within a set of $\mathcal{T}(\mathcal{G}, \mathcal{C})$ with Lebesgue measure zero. Since any finite union of

sets with Lebesgue measure zeros still has Lebesgue measure zero, we conclude that $\lim_{n \rightarrow \infty} \tilde{y}(nT + 1)$ has no identical components generically, which implies that the state of one cluster in $\lim_{n \rightarrow \infty} y(nT + 1)$ are not identical to another generically. This completes the proof.

IV. CLUSTER-CONSENSUS IN DISCRETE-TIME NETWORK WITH SWITCHING TOPOLOGIES

In this section, we study the cluster-consensus in network with switching topologies described as the following time-varying linear system:

$$x_i(t+1) = \sum_{j=1}^N A_{ij}(t)x_j(t) + I_i(t) \quad \forall i \in C_p, \\ p = 1, \dots, K, \quad (14)$$

where $A(t)$ is associated with a graph from the graph set $\Upsilon = \{\mathcal{G}_1, \dots, \mathcal{G}_m\}$ w.r.t. a given clustering \mathcal{C} , each of which satisfy the property \mathcal{A} : for each pair p and q of cluster indices,

- 1) there are not links from C_q to C_p in each graph \mathcal{G}_l , $l = 1, \dots, m$,
- 2) or for each vertex $v \in C_p$ and each graph \mathcal{G}_l , $l = 1, \dots, m$, there are at least one link from C_q to it.

For the matrix sequence $A(t)$, we have the following assumptions:

- \mathcal{B}_1 : There is a positive constant $e > 0$ such that for each pair i, j and t , either $A_{ij}(t) = 0$ or $A_{ij} \geq e$ holds;
- \mathcal{B}_2 : $A_{ii}(t) \geq e$ holds for all $i = 1, \dots, n$ and $t \geq 0$;
- \mathcal{B}_3 (*inter-cluster common influence*): There exists a $\mathbb{R}^{n,n}$ stochastic matrix $B(t) = [b_{p,q}(t)]_{p,q=1}^K$, possibly depending on time, such that

$$\sum_{j \in C_q} A_{ij}(t) = b_{p,q}(t) \quad (15)$$

holds for all $i \in C_p$ and each $p, q = 1, \dots, K$;

- \mathcal{B}_3^* (*static inter-cluster common influence*): There exists a constant $\mathbb{R}^{n,n}$ stochastic matrix $B = [b_{p,q}]_{p,q=1}^K$, possibly depending on time, such that

$$\sum_{j \in C_q} A_{ij}(t) = b_{p,q} \quad (16)$$

holds for all $i \in C_p$ and each $p, q = 1, \dots, K$.

In other words, we define a graph set containing all possible graph induced by the matrix sequence $A(t)$. The graph set satisfies the property in Lemma 5 uniformly and each graph in the set either never occurs in the corresponding graph sequence induced by $A(t)$ or occurs frequently. For each matrix $A(t)$, the nonzero elements should be greater than some constant independent of time and it always has positive diagonals.

Then, we are in the position to give a sufficient condition for the cluster synchronization.

Theorem 3: Suppose that \mathcal{A} , \mathcal{B}_1 , \mathcal{B}_2 and \mathcal{B}_3 hold. If there exists an integer $L > 0$ such that for any L -length time interval $[t, t+L)$, the union graph $\mathcal{G}[\sum_{i=t}^{t+L-1} A(i)]$ has cluster spanning trees, then the system (14) cluster synchronizes.

Proof. The solution of (14) is

$$x(t+1) = A(t)x(t) + \varsigma u(t) = A_0^k x(0) + \sum_{k=0}^t A_{k+1}^t u(k)\varsigma.$$

Noting that the diagonals of each $A(t)$ are positive, we can see that the graph $\mathcal{G}(A_t^{t+L-1})$ contains all links in the graph union $\mathcal{G}(\sum_{k=t}^{t+L-1} A(k))$ and hence has cluster spanning trees and positive diagonals for all t . By Lemma 1, we can conclude that there is an integer N such that the graph $\mathcal{G}(A_t^{t+NL-1})$ is scrambling for all $t \geq 0$. Since the nonzero elements in each $A(t)$ is greater than some constant $e > 0$, there is some $\delta > 0$ such that

$$\inf_t \mu_{\mathcal{C}}(A_t^{t+NL-1}) \geq \delta.$$

Hence, for each t , we have

$$\Delta_{\mathcal{C}}(A_0^t x(0)) \leq (1 - \delta)^{\lfloor \frac{t}{NL} \rfloor} \Delta_{\mathcal{C}}(A(0)),$$

which converges to zero as $t \rightarrow \infty$. Here $\lfloor \cdot \rfloor$ denotes the floor function. Therefore, $\lim_{t \rightarrow \infty} \Delta_{\mathcal{C}}(A_0^t x(0)) = 0$.

Combining with the fact that $\Delta_{\mathcal{C}}(A_t^s \varsigma) = 0$ holds for all $s \geq t$ and ς , we can conclude that the system (14) intra-cluster synchronizes.

The inter-cluster separation can be derived by the same fashion of Theorem 2.

Theorem 4: Suppose that \mathcal{A} , \mathcal{B}_1 , \mathcal{B}_2 and \mathcal{B}_3^* hold. If there exists an integer $L > 0$ such that for any L -length time interval $[t, t+L)$, the union graph $\mathcal{G}[\sum_{i=t}^{t+L-1} A(i)]$ has cluster spanning trees. If the input $u(t)$ and $\sum_{k=0}^t u(k)$ are both bounded, then for any initial data $x(0)$, the solution of (14) is bounded. In addition, if the input $u(t)$ is periodic with a period T and satisfies $\sum_{k=1}^{T-1} u(k) = 0$, (14) reach a cluster consensus generically and each trajectory converges to a T -periodic one.

Proof. To prove the boundedness, we are to find a solution of (14) that stays at $\mathcal{S}_{\mathcal{C}}$ and is the limiting of $x(t)$. Similar to the proof of Theorem 2, we can represent the limiting trajectory by a lower-dimensional linear system (13). The \mathcal{B}_3^* implies that this linear lower-dimensional system is static. So, we can prove its boundedness by the same way of the proof of Theorem 1.

Define the Lyapunov exponent of the matrix sequence $A(t)$ as follows:

$$\lambda(v) = \overline{\lim}_{t \rightarrow \infty} \frac{1}{t} \log \left(\|A_0^t v\| \right).$$

From the Pesin's theory [39], the Lyapunov exponents can only pick finite values and provide a splitting of \mathbb{R}^n . Namely, there is a subspace direct-sum division:

$$\mathbb{R}^n = \oplus_{j=1}^J V_j,$$

and $\lambda_1 > \dots > \lambda_J$, possibly $J < n$, such that for each $v \in V_j$, $\lambda(v) = \lambda_j$. It can be seen that $\lambda_1 = 0$ since $A(t)$, $t \geq 0$, are all stochastic matrices. Let $V = \oplus_{j=1}^J V_j$. From the conditions, we claim

Claim 1: $\mathbb{R}^n = \mathcal{S}_{\mathcal{C}} + V$.

We prove this claim in appendix. Therefore, for any $x(0) \in \mathbb{R}^n$, we can find a vector $y_0 \in \mathcal{S}_{\mathcal{C}}$ such that $x(0) - y_0 \in V$. Define a linear system:

$$y(t+1) = A(t)y(t) + \varsigma u(t), \quad y(0) = y_0. \quad (17)$$

Then, letting $\delta x(t) = x(t) - y(t)$, it should satisfy:

$$\delta x(t+1) = A(t)\delta x(t), \quad \delta x(0) = y(0) - x(0) \in V.$$

Since $\delta x(0) \in V$, $\lambda(\delta x(0)) < 0$. This implies $\lim_{t \rightarrow \infty} \delta x(t) = 0$. So, $\lim_{t \rightarrow \infty} [x(t) - y(t)] = 0$. We can rewrite the equation (17) as a lower-dimensional linear system:

$$\tilde{y}(t+1) = B\tilde{y}(t) + \tilde{\zeta}u(t), \quad (18)$$

which is same with (13). The B_3^* guarantees that the matrix B is static. So, the proof of boundedness of $\tilde{y}(t)$ is an overlap of that of Theorem 1.

In addition, since B is static, then the inter-cluster separation can be proved as an overlap of that of Theorem 1. Therefore, we can conclude that $x(t)$ is bounded, too. This completes the proof.

V. SIMULATION

In this section, we give two numerical examples to illustrate the validity of the proposed theoretic results. We consider graphs with 9 nodes and there are three clusters: $\mathcal{C}_1 = \{1, 2, 3\}$, $\mathcal{C}_2 = \{4, 5, 6\}$ and $\mathcal{C}_3 = \{7, 8, 9\}$. The graph topologies are shown in Fig 1. We simulation cluster consensus in discrete-time networks with the following form:

$$x_i(t+1) = \sum_{q \in \{k | \mathcal{N}_i \cap \mathcal{C}_k \neq \emptyset\}} \frac{\beta_{p,q}}{d_{iq}} \sum_{j \in \mathcal{N}_i \cap \mathcal{C}_q, j \neq i} x_j(t) + \alpha_i u(t), \quad i \in \mathcal{C}_p, p = 1, 2, 3, \quad (19)$$

where $x_i(t) \in \mathbb{R}$, d_{iq} denote the number of agents in set $\mathcal{N}_i \cap \mathcal{C}_q$ and $\{k | \mathcal{N}_i \cap \mathcal{C}_k \neq \emptyset\}$ is identical to all $i \in \mathcal{C}_p, p = 1, 2, 3$. Due to the graph topologies have self-links, we have $d_{ip} \neq 0, \beta_{pp} \neq 0, \forall i \in \mathcal{C}_p$. For any p and $q \in \{k | \mathcal{N}_i \cap \mathcal{C}_k \neq \emptyset\}$, $\sum_{j \in \mathcal{C}_q} \frac{\beta_{p,q}}{d_{iq}} = \sum_{j \in \mathcal{C}_q} \frac{\beta_{p,q}}{d_{i'q}} = \beta_{p,q}$ always holds for $\forall i, i' \in \mathcal{C}_p$, i.e. the coupling matrices to (19) have the common inter-cluster influence. We denote $B = [\beta_{pq}]_{p,q=1}^3$, which inflect the inter-cluster influence among clusters. We choose $u(2j) = -u(2j+1) = 1$, for all $j \in \mathbb{N}$.

We use

$$\eta(x(t)) = \max_{i \in \mathcal{C}_p, j \in \mathcal{C}_q, p \neq q} |x_i(t) - x_j(t)|$$

to measure the difference between clusters and

$$\Delta_C(x(t)) = \max_p \max_{i, i' \in \mathcal{C}_p} |x_i(t) - x_{i'}(t)|$$

to measure the difference of states in the same clusters.

A. Static topology

In this example, the graph is depicted in Fig 1 (a). We take the matrix B of the clustering as:

$$B = \begin{bmatrix} 1 & 0 & 0 \\ 1/3 & 1/3 & 1/3 \\ 0 & 0 & 1 \end{bmatrix}$$

and can see that the graph has cluster spanning trees and the roots of clusters $\mathcal{C}_{1,2,3}$ are 3, 3 and 7 respectively. Therefore, all conditions in Theorem 1 hold. Then (19) reaches cluster consensus generically. The dynamical behaviors of the states $x_i(t), 1 \leq i \leq 9$ are shown in Fig 2 (a) that asymptotically

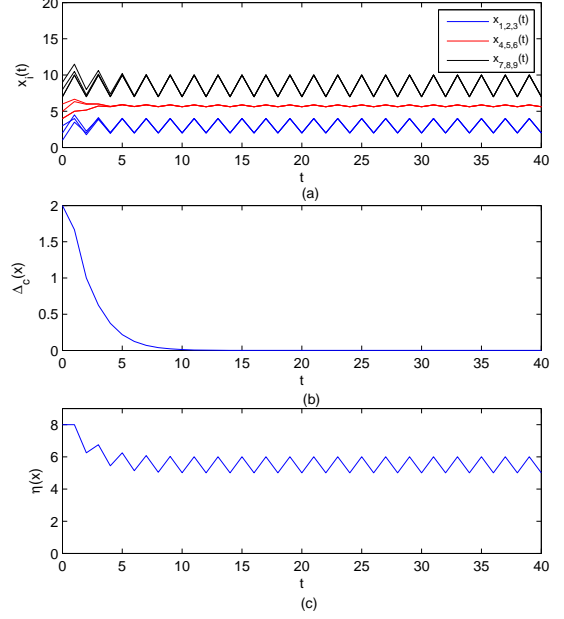


Fig. 2. The dynamical behavior of states $x_i(t)$, $\Delta_C(x(t))$ and $\eta(x(t))$ with respect to example A with randomly picked initial values.

converges to a periodic trajectory with period 2. In Fig 2 (b), the dynamical behaviors of $\Delta_C(x)$ is plotted. It can be seen that $\Delta_C(x)$ converges to zero quickly, which implies that $x(t)$ can intra-cluster synchronize. In Fig 2 (c), the dynamical behaviors of $\eta(x(t))$ is indicated that does not converge to zero but periodically fluctuate. One can see that these states can inter-cluster separate. That is, (19) can reach cluster consensus defined by the three aspects we mentioned.

B. Switching topologies

In this example, the graph topology is switching among the topologies given in Fig 1 (b),(c) and (d) periodically. Noting that none of these graphs has cluster spanning trees, i.e. the condition in Theorem 1 does not hold. However, the union graph of those in Fig 1 (b),(c) and (d) has cluster spanning trees and the roots of clusters $\mathcal{C}_{1,2,3}$ are agents 3, 7 and 7 respectively. We pick an identical matrix B w.r.t. the clustering for the three graphs as

$$B = \begin{bmatrix} 1 & 0 & 0 \\ 1/3 & 1/3 & 1/3 \\ 0 & 0 & 1 \end{bmatrix}.$$

Hence, all assumptions in Theorem 4 holds. Therefore, (19) with switching topologies can achieve cluster consensus. The dynamical behaviors of states $x_i(t), i = 1, \dots, 9$ are shown in Fig 14 (a), the dynamics of the intra-cluster measure $\Delta_C(x)$ is indicated in Fig 14 (b) and the inter-cluster measure $\eta(x(t))$ is plotted in Fig 14(c) respectively. All of them show that the cluster consensus is perfectly reached and $x(t)$ converges to periodic trajectory with period 2.

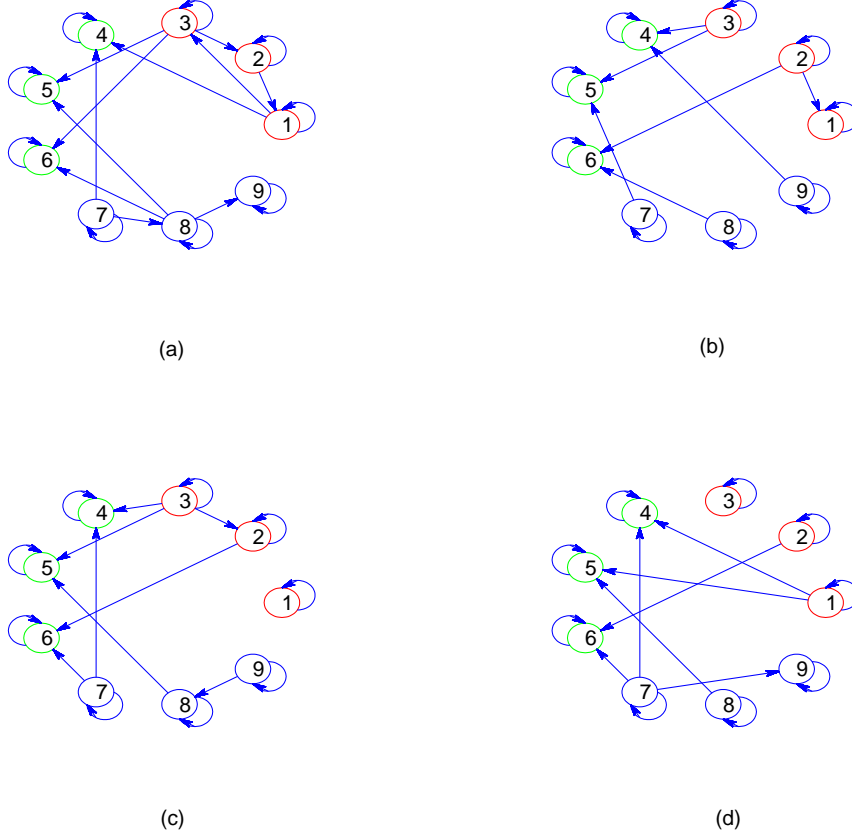


Fig. 1. All of these graphs have self -links. Example A simulate the network with fixed topology (a) and example B simulate the network with topologies switching in (b),(c),(d)

VI. CONCLUSIONS

The idea for studying consensus of multi-agent systems sheds light on cluster consensus analysis. In this paper, we study cluster consensus of multi-agent systems via adaptive inputs. We derive the criterions for cluster consensus in both discrete-time systems with fixed or switching graph topologies. The difference between clustered states are guaranteed by the adaptive inputs of different clusters. We present that if every cluster in the graph corresponding to the system has a spanning tree, then the multi-agent system reaches cluster consensus. The analysis is presented rigorously based on algebraic graph theory and matrix theory. Simple examples are provided to demonstrate the effectiveness of our results.

APPENDIX

In the appendix, we are to prove Claim 1 in the proof of Theorem 3:

$$\mathbb{R}^n = \mathcal{S}_C + V.$$

For this purpose, we define a $\mathbb{R}^{n,n}$ nonsingular matrix $P = [P_1, \dots, P_n]$ with the first K column vectors composing of a

basis of \mathcal{S}_C . In particular, we chose each P_k , $k = 1, \dots, K$, as

$$[P_k]_i = \begin{cases} 1 & i \in \mathcal{C}_k \\ 0 & \text{otherwise.} \end{cases}$$

Then, we can make a linear transformation of the matrices $A(t)$ by:

$$\hat{A}(t) \triangleq P^{-1} A(t) P = \begin{bmatrix} \hat{A}_{1,1} & \hat{A}_{1,2}(t) \\ 0 & \hat{A}_{2,2}(t) \end{bmatrix},$$

where the bottom-left block equals to zero since the subspace \mathcal{S}_C is invariant by $A(t)$ and the top-left block $\hat{A}_{1,1}$ is a static matrix due to B_3^* . Furthermore, since all eigenvalues of B , defined in (15), of which the modules equal to 1 should equal to 1, owing to the fact that all matrices $A(t)$ have all diagonal elements positive, we can select the a new linear transformation Q with the first several columns composing of the basis of the eigenspace of the static sub-matrix $\hat{A}_{1,1}$ corresponding to eigenvalue 1 and all last $n - K$ columns

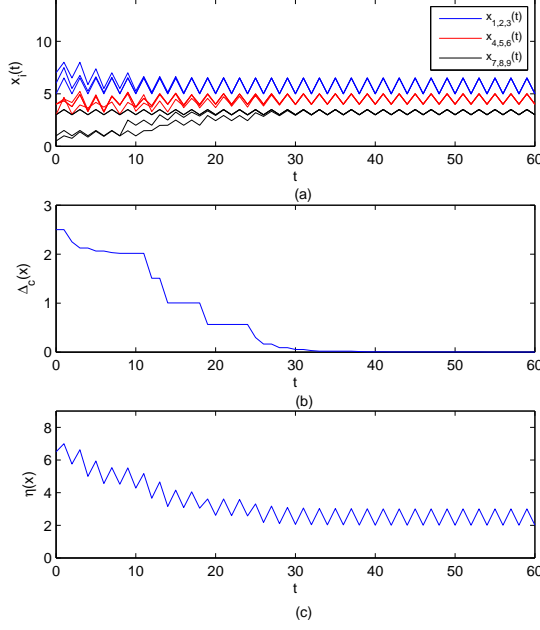


Fig. 3. The dynamical behavior of states $x_i(t)$, $\Delta_c(x(t))$ and $\eta(x(t))$ with respect to example B with randomly picked initial values.

equal to e_j with $j \geq K + 1$, that is, Q has the form as:

$$Q = \begin{bmatrix} Q_1 & 0 \\ 0 & I_{n-K} \end{bmatrix}.$$

Then, we further make linear transformation with Q over $\hat{A}(t)$ resulting in:

$$\tilde{A}(t) \triangleq Q^{-1} \hat{A}(t) Q = \begin{bmatrix} \tilde{A}_{1,1} & \tilde{A}_{1,2}(t) \\ 0 & \tilde{A}_{2,2}(t) \end{bmatrix},$$

where $\tilde{A}_{1,1}$ has the following block form:

$$\tilde{A}_{1,1} = \begin{bmatrix} \tilde{A}_{1,1}^{1,1} & 0 \\ 0 & \tilde{A}_{1,1}^{2,2} \end{bmatrix}.$$

with all eigenvalues of $\tilde{A}_{1,1}^{1,1}$ equal to 1 and $\rho(\tilde{A}_{1,1}^{2,2}) < 1$. Accordingly, we write

$$\tilde{A}_{1,2}(t) = \begin{bmatrix} \tilde{A}_{1,2}^1(t) \\ \tilde{A}_{1,2}^2(t) \end{bmatrix}.$$

Thus, we define

$$\tilde{A}_0^t = \begin{bmatrix} (\tilde{A}_{1,1})^{(t+1)} & \tilde{A}^{(t)} \\ 0 & (\tilde{A}_{2,2})_0^t \end{bmatrix}$$

where $(\cdot)_0^t$ denotes the left matrix product from 0 to t , as defined before.

We define the *projection radius* (w.r.t. \mathcal{C}) of $A(t)$ as follows:

$$\rho_{\mathcal{C}}(A(\cdot)) = \overline{\lim}_{t \rightarrow \infty} \left\{ \|(\tilde{A}_{2,2})_0^{t-1}\| \right\}^{1/t}$$

and the *cluster Hajnal diameter* (w.r.t. \mathcal{C}) of $A(t)$ as follows:

$$\Delta_{\mathcal{C}}(A(\cdot)) = \overline{\lim}_{t \rightarrow \infty} \left\{ \|\Delta_{\mathcal{C}}(A_0^{t-1})\| \right\}^{1/t}$$

for some norm $\|\cdot\|$ that is induced by vector norm. It can be seen that the projection radius and cluster Hajnal diameter are independent of the selection of the matrix norm and the matrix P . First, we shall prove that the projection radius equals to the Hajnal diameter.

Lemma 6: $\rho_{\mathcal{C}}(A(\cdot)) = \Delta_{\mathcal{C}}(A(\cdot))$.

Proof. The proof is quite similar to that in [40] and can be regarded as a generalization of Lemma 2.4 in [40]. For any $d > \rho_{\mathcal{C}}(A(\cdot))$, there exists $T > 0$ such that the inequality

$$\|(\tilde{A}_{2,2})_0^{t-1}\| < d^t$$

for all $t > T$. Then

$$\begin{aligned} & \left\| \tilde{A}_0^{t-1} - \begin{bmatrix} E_K & \\ & 0 \end{bmatrix} \begin{bmatrix} (\tilde{A}_{11})_0^{t-1} & \tilde{A}^{(t-1)} \end{bmatrix} \right\| \\ &= \left\| \begin{bmatrix} 0 & 0 \\ 0 & (\tilde{A}_{2,2})_0^{t-1} \end{bmatrix} \right\| \leq C d^t \end{aligned}$$

for some $C > 0$ and all $t > T$. Thus,

$$\left\| A_0^{t-1} - P \begin{bmatrix} E_K & \\ & 0 \end{bmatrix} \begin{bmatrix} (\tilde{A}_{11})_0^{t-1} & \tilde{A}^{(t-1)} \end{bmatrix} P^{-1} \right\| \leq C_1 d^t,$$

for some $C_1 > 0$ and all $t > T$. Let

$$\begin{aligned} G &= P \begin{bmatrix} E_K & \\ & 0 \end{bmatrix} = [P_1, \dots, P_K], \\ H &= \begin{bmatrix} (\tilde{A}_{11})_0^{t-1} & \tilde{A}^{(t-1)} \end{bmatrix} P^{-1}. \end{aligned}$$

Since each $P_k \in \mathcal{S}_{\mathcal{C}}$ for all $k = 1, \dots, K$, each column vector in the matrix $G \cdot H$ should belong to $\mathcal{S}_{\mathcal{C}}$, too. So, according to the definition of Hajnal diameter, we have

$$\Delta_{\mathcal{C}}(A_0^{t-1}) \leq 2C_1 d^t$$

for all $t \geq T$. This implies that $\Delta_{\mathcal{C}}(A(\cdot)) \leq d$. According to the arbitrariness of d , we have $\Delta_{\mathcal{C}}(A(\cdot)) \leq \rho_{\mathcal{C}}(A(\cdot))$.

On the other hand, for any $d > \Delta_{\mathcal{C}}(A(\cdot))$, there exists $T > 0$ such that $\Delta_{\mathcal{C}}(A_0^{t-1}) < d^t$ holds for all $t > T$. Without loss of generality, we suppose that the clustering \mathcal{C} is successive, i.e., $\mathcal{C}_1 = \{1, 2, \dots, n_1\}$, $\mathcal{C}_2 = \{n_1 + 1, n_1 + 2, \dots, n_2\}, \dots$, $\mathcal{C}_K = \{n_{K-1} + 1, n_{K-1} + 2, \dots, n_K\}$ with $n_K = n$. Select one single row in A_0^{t-1} from each cluster and compose them into a matrix, denoted by H . Let $G = [P_1, \dots, P_K]$. Then the rows of $G \cdot H$ corresponding to the same cluster is identical. So,

$$\|A_0^{t-1} - G \cdot H\| \leq C_2 d^t$$

for some $C_2 > 0$ and $t > T$. Then,

$$\|P^{-1} A_0^{t-1} P - P^{-1} G \cdot H P\| \leq C_3 d^t$$

i.e.,

$$\left\| \begin{bmatrix} (\tilde{A}_{11})_0^{t-1} & \tilde{A}^{(t-1)} \\ 0 & (\tilde{A}_{2,2})_0^{t-1} \end{bmatrix} - \begin{bmatrix} Y & Z \\ 0 & 0 \end{bmatrix} \right\| \leq C_3 d^t$$

for some matrices Y and Z , $C_3 > 0$ and all $t > T$. This implies that $\|(\tilde{A})_0^{t-1}\| \leq C_4 d^t$ holds for some $C_4 > 0$ and all $t > T$. Therefore, $\rho_{\mathcal{C}}(A(\cdot)) \leq d$. The arbitrariness of d can guarantee $\Delta_{\mathcal{C}}(A(\cdot)) \geq \rho_{\mathcal{C}}(A(\cdot))$. From both sides, we have $\Delta_{\mathcal{C}}(A(\cdot)) = \rho_{\mathcal{C}}(A(\cdot))$. This complete the proof of this lemma.

From Theorem 3, we can conclude $\Delta_C(A(\cdot)) < 1$. Thus, $\rho_C(A(\cdot)) < 1$. Thus, for any n -dimensional vector w_0 , we can write it as:

$$w_0 = \begin{bmatrix} z_0 \\ u_0 \\ v_0 \end{bmatrix}$$

where z_0 corresponds to the sub-matrix $\tilde{A}_{1,1}^{1,1}$, u_0 corresponds to the sub-matrix $\tilde{A}_{1,1}^{2,2}$ and $v_0 \in \mathbb{R}^{n-K}$. We rewrite w_0 as a sum of $w_0^1 + w_0^2$ with

$$w_0^1 = \begin{bmatrix} z_0^1 \\ 0 \\ 0 \end{bmatrix}, \quad w_0^2 = \begin{bmatrix} z_0^2 \\ u_0 \\ v_0 \end{bmatrix}$$

where $z_0^1 + z_0^2 = z_0$ that will be determined in the following. It is clear that w_0^1 corresponds a vector in \mathcal{S}_C before the linear transformation PQ . So, if we could pick a suitable z_0^2 such that $\lim_{t \rightarrow \infty} (A)_0^t w_0^2 = 0$, then we would have w_0^2 corresponds a vector in V before QP transformation. This could complete the proof of the claim.

For this purpose, we consider the following linear system:

$$\tilde{w}(t+1) = A(t)\tilde{w}(t), \quad \tilde{w}(0) = w_0^2,$$

which can be rewritten as the following component-wise form:

$$\begin{cases} \tilde{w}_1(t+1) = \tilde{A}_{1,1}^{1,1}\tilde{w}_1(t) + \tilde{A}_{1,2}^1(t)\tilde{w}_3(t) \\ \tilde{w}_2(t+1) = \tilde{A}_{1,1}^{2,2}\tilde{w}_2(t) + \tilde{A}_{1,2}^2(t)\tilde{w}_3(t) \\ \tilde{w}_3(t+1) = \tilde{A}_{2,2}(t)\tilde{w}_3(t) \end{cases}$$

$$\text{with } \tilde{w}_1(0) = z_0^2, \quad \tilde{w}_2(0) = u_0, \quad \tilde{w}_3(0) = v_0.$$

It can be seen that $\lim_{t \rightarrow \infty} \tilde{w}_3(t) = 0$ exponentially because of $\rho_C(A(\cdot)) < 1$ and $\lim_{t \rightarrow \infty} \tilde{w}_2(t) = 0$ exponentially because of $\rho(\tilde{A}_{1,1}^{2,2}) < 1$ and the boundedness of $\tilde{A}_{1,2}(t)$. Without loss of generality, since $\rho_C(A) < 1$ and all eigenvalues of $(\tilde{A}_{1,1}^{1,1})^{-1}$ equal to 1, for any $\epsilon_0 \in (0, |\lambda_2|)$, we have $\|(\tilde{A}_{2,2})_0^t\| \leq M_2 \exp[-(|\lambda_2| - \epsilon_0)t]$, $\|(\tilde{A}_{1,1}^{1,1})^{-1}\| < \exp(\epsilon_0)$ and $\|\tilde{A}_{1,2}^1(t)\| \leq M_0$ for some $M_0 > 0$, $\lambda_0 > 0$, all $t \geq 0$ and some norm $\|\cdot\|$. Note that

$$\tilde{w}_1(t) = (\tilde{A}_{1,1}^{1,1})^t z_0^2 + \sum_{k=0}^t (\tilde{A}_{1,1}^{1,1})^{t-k} \tilde{A}_{1,2}^1(k) [\tilde{A}_{2,2}]_0^k v_0.$$

Since

$$\begin{aligned} & \|(\tilde{A}_{1,1}^{1,1})^{-k}\| \cdot \|\tilde{A}_{1,2}^1(k)\| \|(\tilde{A}_{2,2})_0^k\| \\ & \leq \exp(\epsilon_0 \cdot k/2 - [|\lambda_2| - \epsilon_0] \cdot k) \cdot M_2^2 \\ & \leq \exp(-[|\lambda_2| - 2\epsilon_0]k) M_2^2, \end{aligned}$$

we let

$$R = \sum_{k=0}^{\infty} (\tilde{A}_{1,1}^{1,1})^{-k} \tilde{A}_{1,2}^1(k) [\tilde{A}_{2,2}]_0^k$$

of which the limit exists in the norm sense and the operator R is well-defined. Let consider a subspace of \mathbb{R}^n :

$$\tilde{V} = \left\{ [z^\top, u^\top, v^\top]^\top \in \mathbb{R}^n : z = -Rv \right\}.$$

If we pick z_0^2 such that $w_0^2 \in \tilde{V}$, then we have

$$\begin{aligned} (\tilde{A}_{1,1}^{1,1})^{-t} \tilde{w}_1(t) &= z_0^2 + \sum_{k=0}^t (\tilde{A}_{1,1}^{1,1})^{-k} \tilde{A}_{1,2}^1(k) [\tilde{A}_{2,2}]_0^k v_0 \\ &\rightarrow z_0^2 + Rv_0 = 0 \end{aligned}$$

exponentially as $t \rightarrow \infty$. So, $(\tilde{A})_0^t w_0^2$ converges to zero exponentially. This completes the proof.

REFERENCES

- [1] R. Vidal, O. Shakernia, and S. Sastry, Formation control of non-holonomic mobile robots omnidirectional visual servoing and motion segmentation. *Proc. IEEE Conf. Robotics and Automation*, 2003, pp. 584-589.
- [2] J. Cortes and F. Bullo, Coordination and geometric optimization via distributed dynamical systems. *SIAM J. Control Optim.*, May 2003.
- [3] A. Fax and R. M. Murray, Information flow and cooperative control of vehicle formations. *IEEE Trans. Automat. Contr.*, 49 (2004), pp. 1465-1476.
- [4] C. W. Reynolds, Flocks, herds, and schools: A distributed behavioral model. *Proc. Comp. Graphics ACM SIGGRAPH'87 Conf.*, 1987, vol. 21, pp. 25-34.
- [5] T. Vicsek, A. Czirook, E. Ben-Jacob, I. Cohen, and O. Shochet, Novel type of phase transition in a system of self-derived particles. *Phys. Rev. Lett.*, vol. 75, no. 6, pp. 1226-1229, 1995.
- [6] L. Xiao and S. Boyd, Fast linear iterations for distributed averaging. *Syst. Control Lett.*, vol. 53, pp. 65-78, 2004.
- [7] W. Ren and R. W. Beard, Consensus seeking in multiagent systems under dynamically changing interaction topologies. *IEEE Trans. Automat. Control*, 50(2005), pp. 655-661.
- [8] L. Moreau, Stability of multiagent systems with time-dependent communication links. *IEEE Trans. Automat. Control*, 50(2005), pp. 169-182.
- [9] M. Porfiri and D. J. Stilwell, Consensus seeking over random weighted directed graphs. *IEEE Trans. Automat. Control*, 52(2007), pp. 1767-1773.
- [10] Y. Hatano and M. Mesbahi, Agreement over random networks. *IEEE Trans. Automat. Control*, 50(2005), pp. 1867-1872.
- [11] B. Liu and T. P. Chen, Consensus in networks of multi-agents with cooperation and competition via stochastically switching topologies. *IEEE Trans. Neural Netw.*, 19(2008), pp. 1967-1973.
- [12] R. Olfati-Saber, J. A. Fax and R. M. Murray, Consensus and cooperation in networked multi-agent systems, *Proceedings of the IEEE*, 95(2007), pp. 215-233.
- [13] W. Ren and R. W. Beard, Consensus of information under dynamically changing interaction topologies. *Amer. Control Conf.*, 2004, pp. 4939-4944.
- [14] W. Lu, F. Atay and J. Jost, Consensus and synchronization in discrete-time networks of multi-agents with stochastically switching topologies and time delays. *Networks and Heterogeneous Media*, 6 (2011), pp. 329-349.
- [15] B. Liu, W. Lu and T. Chen, B. Liu, Consensus in networks of multiagents with switching topologies modeled as adapted stochastic processes. *SIAM J. Control Optim.* 49:1 (2011), pp. 227-253.
- [16] J. Hajnal, On products of nonnegative matrices. *Math. Proc. Camb. Phil. Soc.* 52 (1976), pp. 521-530.
- [17] J. Hajnal, Weak ergodicity in non-homogeneous Markov chains. *Proc. Camb. Phil. Soc.*, 54(1958), pp. 233-246.
- [18] J. Shen, A geometric approach to ergodic non-homogeneous markov chains. *Wavelet Anal. Multi. Meth., LNPAM*, 212 (2000), 341-366.
- [19] J. Wolfowitz, Products of indecomposable, aperiodic, stochastic matrices. *Proceedings of the American Mathematical Society*, Vol. 14, No. 5 (Oct., 1963), pp. 733-737.
- [20] R. Olfati-Saber and R. M. Murray, Consensus problems in networks of agents with switching topology and time-delays. *IEEE Trans. Autom. Control*, vol. 49, pp. 1520-1533, 2004.
- [21] X. Liu, W. Lu and T. Chen, Consensus of multi-agent systems with unbounded time-varying delays. *IEEE Transactions on Automatic Control*, 55(2010), pp. 2396-2401.
- [22] X. Liu and T. Chen, Consensus problems in networks of agents under nonlinear protocols with directed interaction topology. *Physical Letters A*, 373 (2009), pp. 3122-3127.
- [23] A. Schnitzler and J. Gross, Normal and pathological oscillatory communication in the brain. *Nat. Rev. Neurosci.* 6 (2005), pp. 285-296.

- [24] K. M. Passino, Biomimicry of bacterial foraging for distributed optimization and control. *IEEE Control Syst. Mag.* 22:3(2002), pp. 52-67.
- [25] E. Montbrio, J. Kurths and B. Blasius, Synchronization of two interacting populations of oscillators. *Phys. Rev. E*, 70(2004), 056125.
- [26] N. F. Rulkov, Images of synchronized chaos: Experiments with circuits. *Chaos*, 6:3(1996), pp. 262-279.
- [27] L. Stone, R. Olinky, B. Blasius, A. Huppert, and B. Cazelles, Complex Synchronization Phenomena in Ecological Systems. *AIP Conf. Proc.*, 622(2002), pp. 476-488.
- [28] K. Hwang, S. Tan, and C. Chen, Cooperative strategy based on adaptive Q-learning for robot soccer systems. *IEEE Trans. Fuzzy Syst.*, 12(2004), pp.569-.
- [29] J. Yu and L. Wang, Group consensus in multi-agent systems with undirected communication graphs. *Proc. 7th Asian Control Conf.*, 2009, pp. 105-110.
- [30] J. Yu and L. Wang, Group consensus in multi-agent systems with switching topologies. *IEEE Conference on Decision and Control*. 2009, pp. 2652-2657.
- [31] Y. Chen, J. H. Lv, F. L. Han and X. H. Yu, On the cluster consensus of discrete-time multi-agent systems . *Systems and control letters* 60(2011), pp. 517-523.
- [32] C. Godsil and G. Royle, "Algebraic Graph Theory," Springer-Verlag, New York, 2001.
- [33] R. A. Horn and C. R. Johnson, "Matrix analysis," Cambridge University Press, 1985.
- [34] C. W. Wu, Synchronization and convergence of linear dynamics in random directed networks, *IEEE Trans. Autom. Control*, 51 (2006), pp. 1207-1210.
- [35] W. L. Lu, B. Liu. T. P. Chen, Cluster synchronization in networks of coupled nonidentical dynamical systems, *CHAOS*, 20 (2010), 013120.
- [36] C.-T. Lin, Structural controlability. *IEEE Trans. Auto. Control*, 19 (1974), pp. 201-208.
- [37] K. J. Reinschke, "Multivariable Control: A Graph-Theories Approach", Springer-Verlag, 1988
- [38] J. -M. Dion, C. Commault, J. van der Woude, Generic properties and control of linear structured systems. *Automatica*, 39 (2003), pp. 1125-1144.
- [39] L. Barreira and Y. Pesin, "Lyapunov exponents and smooth ergodic theory" University Lecture Series, AMS, Providence, RI, 2001.
- [40] W. Lu, F. M. Atay and J. Jost, Synchronization of discrete-time networks with time-varying couplings. *SIAM J. Math. Anal.*, 39 (2007), 1231-1259.